

DOCUMENT RESUME

ED 312 773

EA 021 431

AUTHOR Bowers, Bruce C.
TITLE Alternatives to Standardized Educational Assessment.
ERIC Digest Series Number EA 40.
INSTITUTION ERIC Clearinghouse on Educational Management, Eugene,
Oreg.
SPONS AGENCY Office of Educational Research and Improvement (ED),
Washington, DC.
PUB DATE 89
CONTRACT OERI-R188062004
NOTE 4p.
AVAILABLE FROM Publication Sales, ERIC Clearinghouse on Educational
Management, University of Oregon, 1787 Agate Street,
Eugene, OR 97403 (\$2.50 prepaid postage and
handling).
PUB TYPE Information Analyses -- ERIC Information Analysis
Products (071)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; *Criterion Referenced Tests;
*Educational Assessment; *Educational Testing;
Elementary Secondary Education; *Evaluation Methods;
Mastery Tests; Norm Referenced Tests; *Student
Evaluation; Student Placement
IDENTIFIERS ERIC Digests

ABSTRACT

Standardized testing of American students provides comparative scores for the placement of individual students, enables students to make appropriate decisions regarding a future course of study, and offers the opportunity to assess the effectiveness of teachers, schools, and school districts. As a result of using standardized test scores for comparative purposes, however, "teaching to the test" is becoming a common practice in our nation's schools. Performance-based assessments offer an alternative to standardized testing by requiring students to demonstrate active skills. Because performance-based assessments are more easily scored using a criterion- rather than a norm-referenced approach, to maintain assessment reliability comparisons must be made with similar assessments made by other teachers in other settings, the adoption of an exemplary system involving regular meetings between teachers responsible for scoring would be necessary. Such a shift in testing methodology sacrifices reliability for the benefits of validity. (9 references) (KM)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED312773



ERIC Clearinghouse
on Educational Management

ERIC Digest Series
Number EA 40 • 1989

ALTERNATIVES TO STANDARDIZED EDUCATIONAL ASSESSMENT

By Bruce C. Bowers

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

EA 021 431



ALTERNATIVES TO STANDARDIZED EDUCATIONAL ASSESSMENT

By Bruce C. Bowers

An American educator who was examining the British educational system once asked a headmaster why so little standardized testing took place in British schools. "My dear fellow," came the reply, "In Britain we are of the belief that, when a child is hungry, he should be fed, not weighed." This anecdote suggests the complementary question: "Why is it that we do so *much* standardized testing in the United States?"

What Are the Main Uses of Standardized Testing in American Public Schools?

Advocates of standardized testing assert that it simply achieves more efficiently and fairly many of the purposes for which grading and other traditional assessment procedures were designed. Even critics of standardized testing acknowledge that it has filled a vacuum. As Grant Wiggins (1989a) puts it, "Mass assessment resulted from legitimate concern about the failure of schools to set clear, justifiable, and consistent standards to which it would hold its graduates and teachers accountable."

Standardized testing is currently used to fulfill (1) the *administrative* function of providing comparative scores for individual students so that placement decisions can be made; (2) the *guidance* function of indicating a student's strengths or weaknesses so that he or she may make appropriate decisions regarding a future course of study; and, more recently, (3) the *accountability* function of using student scores to assess the effectiveness of teachers, schools, and even entire districts (Robinson and Craver 1989).

What Problems Have Arisen as a Result of Widespread Use of Standardized Testing?

The phrase "test-driven curriculum" (Livingston, Castle, and Nations 1989) captures the essence of the major controversy surrounding standardized testing. When test scores are used on a comparative basis not only to determine the educational fate of individual students, but also to assess the relative "quality" of teachers, schools, and school districts, it is no wonder that "teaching to the test" is becoming a common practice in our nation's schools. This would not necessarily be a problem if standardized tests provided a comprehensive, in-depth assessment of the knowledge and skills that indicate mastery of a given subject matter. However, the main purpose of standardized testing is to *sort* large numbers of

students in as efficient a manner as possible. This limited goal, quite naturally, gives rise to short-answer, multiple-choice questions. When tests are constructed in this manner, active skills, such as writing, speaking, acting, drawing, constructing, repairing, or any of a number of other skills that can and should be taught in schools are automatically relegated to a second-class status.

What Alternatives to Standardized Testing Have Been Suggested?

It is reasonable to assume that the demand for test results that can be compared across student populations will remain strong. The critical question is whether such results can be obtained from tests that attempt a more comprehensive assessment of student abilities than the present standardized tests are capable of providing. An ancillary, but equally critical, question is whether such tests are too costly to be widely administered.

Suggested alternatives are based on the concept of a "performance-based" assessment. Depending on the subject matter being tested, the performance may consist of demonstrating any of the active skills mentioned above. For example, in the area of writing, drawing, or any of the "artistic expression" skills, it has been suggested that a "portfolio assessment," involving the ongoing evaluation of a cumulative collection of creative works, is the best approach (Wolf 1989). For subjects that require the organization of facts and theories into an integrated and persuasive whole (for example, sciences and social sciences), an assessment modelled after the oral defense required of doctoral candidates has been suggested (Wiggins 1989a).

A third approach, which might be termed the "problem solving model," can be adapted to almost any knowledge-based discipline. It involves the presentation of a problematic scenario that can be resolved only through the application of certain major principles (theories, formulae) that are central to the discipline under examination (Archbald and Newmann 1988).

Can Performance-Based Assessments Be Used to Compare Students Across Different Settings?

Performance-based assessment is more easily scored using a criterion-referenced, rather than a norm-referenced approach. Instead of placing a

student's score along a normal distribution of scores from students all taking the same test, a criterion-referenced approach focuses on whether a student's performance meets a criterion level, normally reflecting mastery of the skills being tested.

How can such an assessment be reliably compared to similar assessments made by other teachers in other settings? It has been suggested that American educators adopt the "exemplary system" being called for in Great Britain. In this system, teachers involved in scoring meet regularly "to compare and balance results on their own and national tests" (Wiggins 1989b), thus increasing reliability across settings. Clearly, however, such an approach (similar to the approach currently in use for the scoring of Advanced Placement essay exams) could be prohibitively expensive if carried out on a large scale. A key question is whether the costs associated with this labor intensive scoring system would be offset by the presumed instructional gains obtained from an assessment model that rewarded a more thorough and holistic approach to instruction.

Have There Been Any Statewide Efforts to Provide Alternatives to Standardized Testing?

California has probably made the greatest effort in this direction, beginning in 1987 with its statewide writing test and continuing with its current development of performance-based assessment in science and history (Massey 1989). The Connecticut Assessment of Educational Progress Program uses a variety of performance tasks in its assessment of science, foreign languages, and business education (Baron 1989). (However, this assessment includes only a sample of students at any given grade level, and, in addition, every year there is change in the subjects for which performance tasks are required.) Vermont education officials are currently seeking legislative approval for funds to pursue a portfolio assessment approach in addition to the current standardized testing (Massey 1989).

What Is the Prognosis for a General Shift Away from Standardized Testing and Toward Performance-Based Testing?

In psychometric terms, the tradeoff in such a shift is to sacrifice reliability for validity. That is, performance-based tests do not lend themselves to a cost- and time-efficient method of scoring that, in addition, provides reliable results. On the other hand, they actually test what the educational system is

presumably responsible for teaching, namely, the skills prerequisite for performing in the real world. The additional costs involved in producing reliable results across different settings for performance-based tests are unknown.

The question is whether a majority of educators will echo the sentiments of George Madaus, director of the Center for the Study of Testing, Evaluation, and Educational Policy, who believes that performance-based testing "is not efficient; it's expensive; it doesn't lend itself to mass testing with quick turnaround time—but it's the way to go" (Brandt 1989).

RESOURCES

- Archbald, Doug A., and Fred M. Newmann. *BEYOND STANDARDIZED TESTING: ASSESSING AUTHENTIC ACADEMIC ACHIEVEMENT IN THE SECONDARY SCHOOL*. Reston, VA: National Association of Secondary School Principals, 1988. 65 pages. ED 301 587.
- Baron, Joan B. "Performance Testing in Connecticut." *EDUCATIONAL LEADERSHIP* 46, 7 (April 1989): 8. EJ 387 136.
- Brandt, Ron. "On Misuse of Testing: A Conversation with George Madaus." *EDUCATIONAL LEADERSHIP* 46, 7 (April 1989): 26-29. EJ 387 140.
- Livingston, Carol, Sharon Castle, and Jimmy Nations. "Testing and Curriculum Reform: One School's Experience." *EDUCATIONAL LEADERSHIP* 46, 7 (April 1989): 23-25. EJ 387 139.
- Massey, Mary. "States Move to Improve Assessment Picture." *ACSD UPDATE* 31, 2 (March 1989): 7.
- Ralph, John, and M. Christine Dwyer. *MAKING THE CASE: EVIDENCE OF PROGRAM EFFECTIVENESS IN SCHOOLS AND CLASSROOMS*. Washington, D.C.: U.S. Department of Education, Office of Educational Research and Improvement, November 1988. 54 pages.
- Robinson, Glen E., and James M. Craver. *ASSESSING AND GRADING STUDENT ACHIEVEMENT*. Arlington, VA: Educational Research Service, 1989. 198 pages.
- Wiggins, Grant. "A True Test: Toward More Authentic and Equitable Assessment." *PHI DELTA KAPPAN* 70, 9 (May 1989a): 703-13. EJ 388 723.
- _____. "Teaching to the (Authentic) Test." *EDUCATIONAL LEADERSHIP* 46, 7 (April 1989b): 41-47.
- Wolf, Dennie P. "Portfolio Assessment: Sampling Student Work." *EDUCATIONAL LEADERSHIP* 46, 7 (April 1989): 35-39. EJ 387 143.

A Product of the ERIC Clearinghouse on Educational Management
College of Education, University of Oregon
Eugene, Oregon 97403